AD_____

Award Number: DAMD17-00-1-0213

TITLE: Estrogen Metabolism Genes and Breast Cancer: An Innovative Approach

PRINCIPAL INVESTIGATOR: Habibul Ahsan, M.D.

CONTRACTING ORGANIZATION: Columbia University
New York, New York 10032

REPORT DATE: July 2004

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2004	Final (1 Jul 2000 - 30 Jun 2004)

4. TITLE AND SUBTITLE
Estrogen Metabolism Genes and Breast Cancer: An Innovative Approach

5. FUNDING NUMBERS
DAMD17-00-1-0213

6. AUTHOR(S)
Habibul Ahsan, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Columbia University
New York, New York 10032

E-Mail: Ha37@columbia.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

We proposed to examine associations of *CYP17*, *CYP19*, *CYP1B1*, and *COMT* polymorphisms and breast cancer risk using an innovative family-based design among families participating in the Metropolitan New York Breast Cancer Registry (MNYBCR), one of the six sites for National Cancer Institute (NCI)-initiated Cooperative Family Registry for Breast Cancer Studies (CFRBCS) project. To date, we have accomplished all tasks in the approved Statement of Work as planned. We found the *CYP19* allele with 11 TTTA repeats to be associated with breast cancer risk in these families. We also found that maternal (but not paternal) carrier status of *CYP19* alleles with 11 repeats tended to be associated with breast cancer risk in daughters (independent of the daughters' own genotype), suggesting a possible in-utero effect of *CYP19*. We found no association of a woman's breast cancer risk with either her own or her mother's genotype of *CYP17*, *CYP1B1*, or *COMT* gene. To date, all the work has been completed and we have published two manuscripts to report the findings.

14. SUBJECT TERMS
No subject terms provided.

15. NUMBER OF PAGES
37

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION:

There is increasing evidence that genes encoding enzymes that are involved in the metabolism of endogenous or exogenous carcinogens play a role in the etiology of many cancers, including breast cancer. The estrogen-metabolism genes that have so far shown promising results include: cytochrome P-450 17 (CYP17), catalyzing the early steps of the estrogen synthesis pathway [1-3], cytochrome P-450 19 aromatase (CYP19), catalyzing the terminal step in estrogen synthesis pathway [4-6], cytochrome P-450 1A1 (CYP1A1), catalyzing the step in the formation of 2-OH estrogen metabolites, cytochrome P-450 1B1 (CYP1B1), catalyzing the step in the formation of 4-OH estrogen metabolites [7-10], and catechol-O-methyl transferase (COMT), catalyzing the excretion of carcinogenic 4-OH estrogen metabolites)[11,12]. To date, all studies examining associations between variants of these estrogen-metabolism genes and breast cancer used a classical case-control design which is susceptible to population stratification bias, and produced inconsistent results. We proposed to examine these associations using an innovative family-based design which avoids such biases, among families participating in the Metropolitan New York Breast Cancer Registry (MNYBCR), one of the six sites for National Cancer Institute (NCI)-initiated Cooperative Family Registry for Breast Cancer Studies (CFRBCS) project.

BODY:

This is the final report for the total duration of the study. Since 1995, MNYR assembled 1,150 families with more than 3,500 individuals. Using the existing MNYR resources and the newly-developed statistical method of Whittemore and Tu, we proposed to test the hypotheses: 1) whether a woman's carrier status of the hypothesized alleles is associated with her breast cancer status, 2) whether parental carrier status of the hypothesized alleles is associated with their daughter's breast cancer risk. To date, all tasks listed in the Statement of Work have been accomplished. Detailed reports of findings obtained to date are included in two publications [13,14].

Table 1 presents the FS and NFS for testing the associations between the *a priori* hypothesized variant alleles and breast cancer. Each test statistic has approximately a standard Gaussian distribution under the null hypothesis of no association between genotype and breast cancer risk. A positive value of a NFS reflects excess transmission of the variant allele to affected daughters, and a negative value represents fewer such transmissions than expected under the null. Thus a test statistic that is negative but large in absolute value would suggest that the variant allele is associated with reduced risk. We computed the FS and NFS under recessive, dominant and additive models. For the dominant models, the number of affected daughters carrying one or more copies of the variant alleles was compared to that expected from the parental genotypes according to Mendelian expectation. Similarly, for the recessive models, the number of affected daughters homozygous for the variant allele is compared to that expected under

Mendelian expectation. For the additive models, the total variant allele count in affected daughters is compared to that expected from the parental genotypes according to Mendelian expectation.

As seen in Table 1, the NFS for association between *CYP19* $TTTA_{11}$ allele and breast cancer under the dominant model is 1.83, which is higher than the critical value (1.65) for a one-tailed test statistic, suggesting that affected daughters were more likely to receive the *CYP19* $TTTA_{11}$ allele from their parents (irrespective of their ethnic distribution) than unaffected daughters. Like the NFS, the FS was also statistically significant under the dominant model, supporting an association between the *CYP19* $TTTA_{11}$ allele and breast cancer among the parents in these families. The results for *CYP19* $TTTA_{\geq 10}$ alleles did not show a consistent association, as only the FS was statistically significant under dominant model. Neither the FS nor the NFS suggested any significant association between the *CYP17* variant C allele and breast cancer, under any of the models of FGAP analyses. The FSs for the association between the *CYP1B1* codon 453 variant C allele and breast cancer were statistically significant (2.33 and 1.77 under the recessive and additive models, respectively, which are higher than the critical value 1.65) for a one-tailed test statistic. Similarly, the *FS* for the association between the *COMT* codon 158 variant A allele and breast cancer was statistically significant (1.81 under the additive model, which is higher than the critical value 1.65) for a one-tailed test statistic. However, in the absence of any association based on *NFS* (which is more robust) the associations based on *FS* are unlikely to be meaningful.

Table 1. Association between variant alleles of estrogen-metabolism genes and breast cancer

Allele Type	Association Between CYP17/CYP19 Variant Alleles and Breast Cancer						
	Estimated population allele frequency (%)	Nonfounder Statistic			Founder Statistic		
		Recessive Model	Dominant Model	Additive Model	Recessive Model	Dominant Model	Additive Model
CYP1B1 codon 432							
C	42.77	0.32	-0.83	-0.43	1.35	0.44	1.12
(p-value)		(0.38)	(0.20)	(0.33)	(0.09)	(0.33)	(0.13)
CYP1B1 codon 453							
G	16.09	0.59	0.43	0.64	**2.33**	1.09	**1.77**
(p-value)		(0.28)	(0.33)	(0.26)	(0.01)	(0.14)	(0.04)
COMT							
A	49.40	-0.48	-1.01	-1.08	1.47	1.30	**1.81**
(p-value)		(0.32)	(0.16)	(0.14)	(0.07)	(0.10)	(0.04)
CYP17							
C	42.46	-1.01	-1.52	-1.85	0.40	1.08	1.01
(p-value)[1]		(0.16)	(0.06)	(0.03)	(0.34)	(0.14)	(0.16)
CYP19 (# of TTTA repeats)							
$(TTTA)_{\geq 10}$	33.71	-1.24	1.26	0.32	-0.32	**1.66**	1.13
(p-value)[1]		(0.11)	(0.10)	(0.38)	(0.37)	(0.05)	(0.13)
$(TTTA)_{11}$	28.78	-1.09	**1.83**	0.97	-1.50	**1.96**	0.89
(p-value)[1]		(0.14)	(0.03)	(0.17)	(0.07)	(0.03)	(0.19)

Table 2 presents results relating maternal and paternal carrier statuses for the variants of estrogen-biosynthesis genes with breast cancer risk in daughters. Mothers of affected daughters were more likely to carry the *CYP19* alleles with 11 TTTA repeats ($TTTA_{11}$) than expected in the parental population. There were no such associations between the paternal carrier status of $TTTA_{11}$ or any of the other *CYP19* alleles and breast cancer in daughters. For this hypothesis, the findings for analysis involving *CYP19* $TTTA_{\geq 10}$ corroborated with that of $TTTA_{11}$ alleles. Although maternal carrier status of the *CYP17* C allele tended to be positively associated with daughter's breast cancer this association was not specific to the mothers but was present also among the fathers. Mothers (and fathers) of higher numbers of affected daughters were more likely to carry the *CYP1B1* codon 453 G allele and the *COMT* codon 158 variant A allele than mothers (and fathers) of fewer unaffected daughters. Although parental carrier status of the *CYP1B1* codon 432 C allele tended to be positively associated with daughter's breast cancer this association was not statistically significant.

In conclusion, we found the *CYP19* allele with 11 TTTA repeats to be associated with breast cancer risk in these families. We also found that maternal (but not paternal) carrier status of *CYP19* alleles with 11 repeats tended to be associated with breast cancer risk in daughters (independent of the daughters' own genotype), suggesting a possible in-utero effect of *CYP19* [13]. We found no association of a woman's breast cancer risk with either her own or her mother's genotype of *CYP17, CYP1B1*, or *COMT* gene [14]. This was consistent with results from matched case-control analyses using all available sibships in these families. We also found that parental carrier status of the *CYP1B1* codon 453 variant G allele and the *COMT* codon 158 variant A allele was associated with breast cancer risk in daughters [14].

Table 2. Association between mothers' and fathers' carrier status of the variant allele(s) and breast cancer risk in daughters

Variant Allele(s)	Mothers' Carrier Status	Fathers' Carrier Status
CYP1B1 codon 432		
C	1.47	1.58
(p-value)	(0.07)	(0.06)
CYP1B1 codon 453		
G	1.95	3.17
(p-value)	(0.03)	(<0.001)
COMT		
A	1.90	2.45
(p-value)	(0.03)	(0.01)
CYP17		
C	1.47	1.40
(p-value)[2]	(0.07)	(0.08)
CYP19 (#of repeats)		
$(TTTA)_{\geq 10}$	**1.73**	0.95
(p-value)[2]	(0.04)	(0.17)
$(TTTA)_{11}$	1.52	-0.11
(p-value)[2]	(0.06)	(0.46)

*FS was calculated under additive model

KEY RESEARCH ACCOMPLISHMENTS:

Task 1. Identify study subjects and their extracted DNA (Months 1-5)

1. Identify breast cancer cases in the registry who provided blood and have at least one member of the parents who also provided blood.
   Track the DNA aliquots stored in Dr. Santella=s laboratory for these subjects.

Task 2. Determine CYP17, CYP19, COMT and CYP1B1 genotype of 150 cases and their parents (Months 6-24)

2a. PCR amplification and RFLP assay on 450 cases and controls.
2b. Track laboratory procedures on personal computer

Task 3. Data entry and statistical analysis (Months 25-36)

3a. Enter laboratory data into personal computer
3b. Merge laboratory data with the interview data
3c. Conduct statistical analysis examining the association of breast cancer in relation to CYP17, CYP19, COMT and CYP1B1 genotype
3d. Examine the associations with respect to BRCA1/2, the reproductive and hormonal risk factor status of the index case
3e. Manuscript preparation

To date, all the major tasks listed in the Statement of Work have been accomplished, including identification of study subjects and their extracted DNA, determination of *CYP17, CYP19, COMT* and *CYP1B1* genotype of 150 cases and their parents, and Data entry and statistical analysis. Two manuscripts have been prepared and submitted for publication. We were unable to examine the associations with respect to BRCA1/2 status (Task 3d) given the very small number of carriers with mutations in the BRCA1/2 genes tested to date. This component was dependent on the completion of BRCA1/2 mutation testing of the study participants as part of the parent MNYR project (not this DOD project) and the mutation testing of all families are yet to be fully completed.

REPORTABLE OUTCOMES:

Publications:

1) Variants in Estrogen-biosynthesis Genes *CYP17* and *CYP19* and Breast Cancer Risk: A Family-based Genetic Association Study. Ahsan H, Whittemore AS, Senie RT, Hamilton SP, Wang Q, Gurvich I, Santella RM. *Breast Cancer Research*. In Press

2) A family-based genetic association study of variants in estrogen-metabolism genes COMT and CYP1B1 and breast cancer risk. Ahsan H, Chen Y, Whittemore AS, Kibriya MG, Gurvich I, Senie RT, Santella RM. *Breast Cancer Res Treat.* 2004 May;85(2):121-31.

Abstracts and presentations:

1) Ahsan H, Chen Y, Whittemore AS, Senie RT, Kibriya MG, Wang Q, Santella RM. Variants in estrogen-metabolism genes and breast cancer. Presented at *the third Era of Hope meeting for the Department of Defense (DOD) Breast Cancer Research Program (BCRP)*. Orlando, Florida, September 25-28, 2002.

2) Ahsan H, Chen Y, Whittemore AS, Senie RT, Kibriya MG, Wang Q, Santella RM. Variants in estrogen-metabolism genes and breast cancer. Presented at the *Molecular and Genetic Epidemiology of Cancer, An AACR International Conference co-sponsored by the Society of Toxicology in conjunction with the Molecular Epidemiology Working Group of the AACR*. Waikoloa, Hawaii, January 18-23

CONCLUSIONS:

We have accomplished all the major tasks in the approved Statement of Work as planned. Two manuscripts have been published.

We found the *CYP19* allele with 11 TTTA repeats to be associated with breast cancer risk in these families. We also found that maternal (but not paternal) carrier status of *CYP19* alleles with 11 repeats tended to be associated with breast cancer risk in daughters (independent of the daughters' own genotype), suggesting a possible in-utero effect of *CYP19*. We found no association of a woman's breast cancer risk with either her own or her mother's genotype of *CYP17, CYP1B1*, or *COMT* gene. We also found that parental carrier status of the *CYP1B1* codon 453 variant G allele and the *COMT* codon 158 variant A allele was associated with breast cancer risk in daughters. Since this is the first study to report an association with the *CYP19* $TTTA_{11}$ allele, and since multiple comparisons have been made, the associations reported in this study should be interpreted with caution and need to be confirmed in future studies.

Although the study found that parental carrier status of certain *CYP1B1* or *COMT* genotypes might be associated with daughter's breast cancer risk, the biological basis as well as independent confirmation of this finding need to be investigated in future larger family-based studies before making meaningful inferences. While these novel associations, if real, may prove to have important implications in advancing our understanding of the breast cancer etiology, we believe that future studies need to confirm them before any meaningful inferences can be made.

REFERENCES:

[1] K.Nedelcheva, V, E.K.Haraldsen, K.B.Anderson, P.E.Lonning, B.Erikstein, R.Karesen, O.S.Gabrielsen, A.L.Borresen-Dale, CYP17 and breast cancer risk: the polymorphism in the 5' flanking area of the gene does not influence binding to Sp-1. Cancer Res. 59 (1999) 2825-2828.

[2] H.S.Feigelson, L.S.Shames, M.C.Pike, G.A.Coetzee, F.Z.Stanczyk, B.E.Henderson, Cytochrome P450c17alpha gene (CYP17) polymorphism is associated with serum estrogen and progesterone concentrations. Cancer Res. 58 (1998) 585-587.

[3] C.A.Haiman, S.E.Hankinson, D.Spiegelman, G.A.Colditz, W.C.Willett, F.E.Speizer, K.T.Kelsey, D.J.Hunter, The relationship between a polymorphism in CYP17 with plasma hormone levels and breast cancer. Cancer Res. 59 (1999) 1015-1020.

[4] C.A.Haiman, S.E.Hankinson, D.Spiegelman, V.De, I, G.A.Colditz, W.C.Willett, F.E.Speizer, D.J.Hunter, A tetranucleotide repeat polymorphism in CYP19 and breast cancer risk. Int.J.Cancer 87 (2000) 204-210.

[5] S.W.Baxter, D.Y.Choong, D.M.Eccles, I.G.Campbell, Polymorphic variation in CYP19 and the risk of breast cancer. Carcinogenesis 22 (2001) 347-349.

[6] Y.Miyoshi, K.Iwao, N.Ikeda, C.Egawa, S.Noguchi, Breast cancer risk associated with polymorphism in CYP19 in Japanese women. Int.J.Cancer 89 (2000) 325-328.

[7] L.R.Bailey, N.Roodi, W.D.Dupont, F.F.Parl, Association of cytochrome P450 1B1 (CYP1B1) polymorphism with steroid receptor status in breast cancer. Cancer Res. 58 (1998) 5038-5041.

[8] W.Zheng, D.W.Xie, F.Jin, J.R.Cheng, Q.Dai, W.Q.Wen, X.O.Shu, Y.T.Gao, Genetic polymorphism of cytochrome P450-1B1 and risk of breast cancer. Cancer Epidemiol Biomarkers Prev. 9 (2000) 147-150.

[9] J.Watanabe, T.Shimada, E.M.Gillam, T.Ikuta, K.Suemasu, Y.Higashi, O.Gotoh, K.Kawajiri, Association of CYP1B1 genetic polymorphism with incidence to breast and lung cancer. Pharmacogenetics 10 (2000) 25-33.

[10] N.A.Kocabas, S.Sardas, S.Cholerton, A.K.Daly, A.E.Karakaya, Cytochrome P450 CYP1B1 and catechol O-methyltransferase (COMT) genetic polymorphisms and breast cancer susceptibility in a Turkish population. Arch.Toxicol. 76 (2002) 643-649.

[11] K.Mitrunen, N.Jourenkova, V.Kataja, M.Eskelinen, V.M.Kosma, S.Benhamou, D.Kang, H.Vainio, M.Uusitupa, A.Hirvonen, Polymorphic catechol-O-methyltransferase gene and breast cancer risk. Cancer Epidemiol Biomarkers Prev. 10 (2001) 635-640.

[12] M.Bergman-Jungestrom, S.Wingren, Catechol-O-Methyltransferase (COMT) gene polymorphism and breast cancer risk in young women. Br.J Cancer 85 (2001) 859-862.

[13] H.Ahsan, A.S.Whittemore, Y.Chen, R.T.Senie, S.P.Hamilton, Q.Wang, I.Gurvich, and R.M.Santella. Variants in estrogen-biosynthesis genes *CYP17* and *CYP19* and breast cancer risk: A family-based genetic association study. Breast Cancer Research. In press

[14] H.Ahsan, Y.Chen, A.S.Whittemore, M.Kibriya, I.Gurvich, R.T.Senie, R.M.Santella, Family-based genetic association study of variants in estrogen-metabolism genes *COMT* and *CYP1B1* and breast cancer risk. Breast Cancer Res Treat. 2004 May;85(2):121-31.

APPENDICES:

Publications:

1) Variants in Estrogen-biosynthesis Genes *CYP17* and *CYP19* and Breast Cancer Risk: A Family-basedGenetic Association Study. Ahsan H, Whittemore AS, Senie RT, Hamilton SP, Wang Q, Gurvich I, Santella RM. *Breast Cancer Research*. In Press

2) A family-based genetic association study of variants in estrogen-metabolism genes COMT and CYP1B1 and breast cancer risk. Ahsan H, Chen Y, Whittemore AS, Kibriya MG, Gurvich I, Senie RT, Santella RM. *Breast Cancer Res Treat*. 2004 May;85(2):121-31.

Appendix 1

Variants in Estrogen-biosynthesis Genes *CYP17* and *CYP19* and Breast Cancer Risk: A Family-based Genetic Association Study*

Habibul Ahsan[1,2,6], MD, MMedSc
Alice S. Whittemore[5], PhD
Yu Chen[1], MPH
Ruby T. Senie[1,2], PhD
Steven P. Hamilton[4], MD, PhD
Qiao Wang[2]
Irina Gurvich[2], MS
Regina M. Santella[2,3], PhD

[1]*Department of Epidemiology, Mailman School of Public Health, Columbia University*
[2]*Herbert Irving Comprehensive Cancer Center, Columbia University*
[3]*Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University*
[4]*Department of Psychiatry, Columbia University College of Physicians and Surgeons and the New York State Psychiatric Institute*
[5]*Division of Epidemiology, Department of Health Research and Policy, Stanford University School of Medicine*

[6]Correspondence and reprint requests should be addressed to: Dr. Habibul Ahsan, Department of Epidemiology, Mailman School of Public Health, 722 West 168th Street, Room 720G, New York, N.Y. 10032. Phone: (212) 305 4657; Fax: (212) 305 9413; E-mail: ha37@columbia.edu.

* This research was supported by U.S. Department of Defense Grant # DAMD17-00-1-0213 (to HA) and National Cancer Institute under RFA # CA-95-003, Breast Cancer Family Registry, through cooperative agreement with Columbia University (to RTS).

Abstract

Case-control studies have reported inconsistent results concerning breast cancer risk and polymorphisms in genes that control endogenous estrogen biosynthesis. We report findings from the first family-based association study examining associations between female breast cancer risk and polymorphisms in two key estrogen-biosynthesis genes *CYP17* (T->C promoter polymorphism) and *CYP19* (TTTA repeat polymorphism). We conducted the study among 278 nuclear families containing one or more daughters with breast cancer, with a total of 1,123 family members (702 with available constitutional DNA and questionnaire data and 421 without them). These nuclear families were selected from breast cancer families participating in the Metropolitan New York Registry (MNYR) - one of the six centers of NCI's Cooperative Family Registry for Breast Cancer Studies. We used likelihood-based statistical methods to examine allelic associations. We found the *CYP19* allele with 11 TTTA repeats to be associated with breast cancer risk in these families. We also found that maternal (but not paternal) carrier status of *CYP19* alleles with 11 repeats tended to be associated with breast cancer risk in daughters (independent of the daughters' own genotype), suggesting a possible in-utero effect of *CYP19*. We found no association of a woman's breast cancer risk with either her own or her mother's *CYP17* genotype. Findings from this family-based study indicate that a woman's personal and maternal carrier status of *CYP19* 11 TTTA repeat allele may be related to increased breast cancer risk. However, since this is the first study to report an association between *CYP19* 11 TTTA repeat allele and breast cancer, and since multiple comparisons have been made, the associations should be interpreted with caution and need confirmation in future family-based studies.

Introduction and Background

Cumulative exposure to circulating estrogen is considered to be of primary importance in breast cancer etiology. Estrogen biosynthesis, cellular binding and metabolism involve many steps and the genes controlling these steps may contribute to inherent variability in breast cancer susceptibility. Endogenous estrogen is produced predominantly in the ovarian theca cells in premenopausal women and in the breast stromal adipose cells in postmenopausal women. The current study focuses on *CYP17* and *CYP19*, two key genes that control the biosynthesis of estradiol and estrones from their lipid precursors and that are expressed in these cells. CYP17 controls two successive early steps of endogenous estrogen biosynthesis by converting pregnenolone and progesterone to precursors of androgen and estrogen. CYP19, also known as aromatase, controls the terminal step of estrogen biosynthesis by converting 19-carbon steroids (testosterone and androstenedione) to 18-carbon estrogens (estradiol and estrone).

A T->C single nucleotide polymorphism (SNP) in the 5' promoter region of the *CYP17* gene and a TTTA repeat polymorphism in the exon 4-intron 5 boundary region of the *CYP19* gene have been investigated in breast cancer by a number of studies with inconsistent results (1, 2). For both polymorphisms the variant alleles are considered to be related to an increased biosynthesis of endogenous estrogen. The *CYP17* T->C polymorphism is thought to create an Sp1-type (CCACC) promoter site (although one study did not confirm this (3)) and is associated with an increased serum estrogen level (4, 5). After Feigelson and colleagues first published their study showing a higher risk of breast cancer in relation to the *CYP17* C allele among non-Caucasian women (6) many other authors attempted to replicate this in other populations. While some studies confirmed this initial finding others did not. All studies reporting an increased risk, including the original study, found the increased risk in one or more certain subgroups of women studied, e.g., women with advanced disease (6), women under 40 (7), women under 40 with family history (8), women over 55 years (9), and women who also carry other genetic polymorphisms (10). Two studies found that women carrying *CYP17* C allele are less likely to use hormone replacement therapy (5, 11) and three studies found that the protective effect of later age at menarche is stronger among women who do not carry the C allele (5, 6, 12). A recent meta-analysis concluded that the *CYP17* T->C polymorphism is not a significant independent risk factor for breast cancer (2).

The *CYP19* gene contains a variable number (ranging 7-13) of TTTA repeats in the exon 4-intron 5 boundary region, creating polymorphisms which have been examined in 5 studies (13-17). Kristensen and colleagues (13) and subsequently others found about a 2-4 fold elevated risk in relation to certain numbers of *CYP19* TTTA repeat polymorphisms. Although one small study found a higher risk in relation to the TTTA 7 repeat ($TTTA_7$) allele (14) most studies reporting an association found elevated risks in relation to one of the higher number of TTTA repeats: 12 repeat ($TTTA_{12}$) (13), 10 repeat ($TTTA_{10}$) (15, 16) or $\geq$10 repeat ($TTTA_{\geq 10}$) alleles (17). A meta-analysis published in 1999 based on some of the earlier studies found that women carrying *CYP19* $TTTA_{10}$ allele were at higher risk of breast cancer (1).

All previously published studies on the topic examining the association between the *CYP17* and *CYP19* polymorphisms and breast cancer discussed above used a classical case-control design. Given that case-control studies may be susceptible to population stratification bias it is important to examine these potentially important biologically plausible hypotheses in family-based studies which are free from such bias. In this study, we examine the association between the *CYP17* promoter T–>C and *CYP19* TTTA repeat polymorphisms and female breast cancer using a family-based design among nuclear families participating in the Metropolitan New York

Registry (MNYR) - one of the six international centers of NCI's Cooperative Family Registry for Breast Cancer Studies (CFRBCS) project. Although other polymorphisms in the *CYP17* and *CYP19* genes have been reported we focused on these two polymorphisms since they have been studied most extensively both in relation to their potential associations with breast cancer and also in relation to their influence on circulating estrogens.

All published studies focused on the relationship between a woman's own constitutional genotype and her breast cancer risk. A body of recent literature has provided limited data suggesting that a woman's breast cancer risk may be related not only to her own endogenous estrogens during adolescence and adulthood, but also to her prenatal exposure, i.e., in-utero exposure to her maternal circulating estrogens (18-23). In addition to the main association between a woman's own genotype and her breast cancer status, the family-based design of the current study allows us to address this hypothesis indirectly, by examining the association between maternal carrier status of *CYP17* or *CYP19* gene variants (i.e., in-utero exposure to an altered level of maternal estrogens) and breast cancer status in daughters.

Methods

Selection of Study Participants

Since 1995, the MNYR has been recruiting families with breast and/or ovarian cancers in clinical and community settings within the metropolitan New York area. Families meeting one or more of the following criteria are invited to participate: a female <45 years of age at diagnosis of breast cancer; a female with both breast and ovarian cancer; three or more relatives with breast or ovarian cancer diagnosed at age ≥45 years, or any male with breast cancer. Following identification of a proband he/she is invited to participate in the registry and his/her family's eligibility is assessed. If the family is eligible and the proband agrees to participate, following appropriate informed consent, he/she is interviewed either in-person or by phone with an epidemiology questionnaire and a family history questionnaire. The proband also is asked to provide permission to contact family members. Blood or buccal samples are also collected and participants are provided with a self-administered dietary questionnaire to be returned by mail. Once family members consent to participate, data and blood or buccal samples from the family members are also collected in a similar manner. For members affected with cancer, tumor tissue samples are collected and reviewed pathologically. Genomic DNA from white blood cells or buccal samples has been collected for participants who donated biological samples. To date, the MNYR has enrolled 1,158 families and over 3,900 total participants.

For this study, we restricted attention to nuclear families having at least one affected daughter with DNA and at least one parent and/or sibling with DNA. Of the 1,158 families enrolled in the MNYR to date, 278 families met these eligibility criteria. Subjects can participate in the MNYR with or without completion of the full epidemiology questionnaire and/or blood samples. There were 1,123 family members in the 278 eligible nuclear families (with a total of 1,123 family members) of which 702 completed the full epidemiology questionnaire and provided blood samples. However, accurate data on relevant variables for the statistical method used in this study (see below) for the remaining 421 members were available from the family-history questionnaire completed by the 702 members. There was 99% concordance in data on age and affected status between women who completed the full epidemiology questionnaire and women who did not.

Laboratory Analysis

We evaluated association between the T–>C single nucleotide polymorphism (SNP) in the promoter region of the *CYP17* gene and the tetranucleotide (TTTA) repeat polymorphism in exon 5/intron 4 boundary of the *CYP19* gene. A total of 23 subjects could not be genotyped for *CYP17* and 26 subjects could not be genotyped for *CYP19*. Genotype data were available on a total of 679 members (from 277 nuclear families) for *CYP17* and 676 members (from 278 nuclear families) for *CYP19*.

The *CYP17* promoter polymorphism was determined using template-directed primer extension and detection by fluorescence polarization in a 96-microwell-based format (24, 25). Briefly, DNA isolated from blood cells by salting out was used for genotyping subjects. First, the target DNA was PCR-amplified (using forward primer TTTAAAAGGCCTCCTTGTGC and reverse primer TTGGGCCAAAACAAATAAGC) to generate products in the range of 100-200 base pairs. After PCR amplification, the primers were digested with shrimp alkaline phosphatase and E. coli exonuclease I. Then, single nucleotide extension was carried out in the presence of the appropriate allele specific ddNTPs differentially fluorescence-labeled with either R110 or TAMRA purchased from NEN Life Sciences (Boston MA). For single nucleotide extension reaction both forward and reverse probes were tested to select the optimum (the forward probe GCCACAGCTCTTCTACTCCAC) based on clear signal differences. The incorporation resulted in diminished rotation of the fluor compared to the ddNTP. Finally, the fluorescence polarization was read on a Tecan Polarion fluorescence polarization microplate reader (Research Triangle Park, NC). The reader generates the genotype data based on the distinct separations (with appropriate cut-offs) of the fluorescent intensity values for different alleles as compared to internal controls.

The *CYP19* TTTA repeats were determined by PCR amplification (using forward primer GTCTATGAATGTGCCTTTTT and reverse primer GTTTGACTCCGTGTGTTTGA) followed by analysis on an ABI 377 system with GenScan software based on the separations on gel according to the differences in the number of TTTA repeats.

All laboratory assays were performed with laboratory personnel blinded to the subject's disease status or family relationships. In addition to assay specific quality control samples, 10% of samples were reassayed after relabeling to keep laboratory staff blinded to its identity.

Statistical Analysis

We used the Family Genetic Analysis Program (FGAP (26)) to test the null hypothesis of no association between genotype and breast cancer risk in nuclear families. The FGAP computes two test statistics: 1) the *nonfounder statistic (NFS)*, a generalization of the transmission disequilibrium test (TDT) (27, 28) which evaluates transmission disequilibrium from parents to offspring, and 2) the *founder statistic (FS)*, which compares the distribution of parental genotypes to that expected under the null hypothesis of no association. The FGAP statistics fully exploit data from families with variable numbers of affected/unaffected members with variable (known/unknown) patterns of parental genotypes.Based on the prior evidence, (6)' (13) (15-17)we hypothesized that breast cancer risk is elevated among carriers of the *CYP17* C allele and the *CYP19* variant alleles with 10 or more TTTA repeats (i.e., the $TTTA_{10}$, $TTTA_{11}$, $TTTA_{12}$, and $TTTA_{13}$ alleles). The data analysis was focused on two specific components of the study hypotheses: 1) whether a woman's carrier status of the hypothesized alleles is associated with her breast cancer status, 2) whether a mother's carrier status of the hypothesized alleles is associated with her daughter's breast cancer risk. For testing the first component of a hypothesis, we applied

the FS and NFS to assess whether specific genotypes of each of the studied genes are related to breast cancer. Since FS and NFS follow a normal Gaussian distribution under the null hypothesis, the assessment of statistical significance of the association can be done based on the deviation of these statistics from the standard critical values under normal distribution.

For simplicity, we describe these analyses for the *CYP17* gene as applied to nuclear families consisting of two parents and at least one daughter. Parents may be untyped and the mother's breast cancer status may be unknown. The test statistics, which are likelihood-based score statistics, are obtained by summing the score contributions from each family. These family-specific scores are obtained in three steps.

In the first step, we impute a probability distribution for the genotypes of each pair of parents, conditional on the observed genotypes of all family members. To do this, we obtain maximum likelihood estimates of the genotypes TT, TC and CC for each of a pair of parents, given the observed genotypes in the family. These estimates do not require the assumption of Hardy-Weinberg frequencies for parental genotypes. If, for example, both parents' genotypes were known, then the probabilities are degenerate at the observed genotypes. Similarly, if both parents' genotypes were unknown but two offspring had observed *CYP17* genotypes TT and CC, then the parental distributions are degenerate at TC since both parents must be heterozygote.

In the second step, we used the inferred parental genotype distribution and the offspring's observed genotypes to evaluate disequilibrium in the transmission of *CYP17* C from parent to daughter. We looked specifically at disequilibrium in which affected daughters received more copies of *CYP17* C than do unaffected daughters. This evaluation was done via the NFS. Under the null hypothesis of no transmission disequilibrium from parent to affected and unaffected offspring, the NFS has an asymptotic standard Gaussian distribution. The NFS generalizes the transmission disequilibrium test (TDT) to families with untyped parents and to families with both affected and unaffected daughters. It can be considerably more powerful than the sibling TDT (S-TDT) test (29) when applied to families without unaffected daughters.

In the final step, we used the inferred parental genotypes (and the mothers' breast cancer phenotypes) in the FS to compare the parental genotype distribution to the expected distribution under the null hypothesis of no association. This statistic treats the affected and unaffected mothers like cases and controls in a case-control study. However, each parent's contribution is weighted in proportion to his/her number of affected and unaffected daughters, so that parents of many affected daughters receive higher weights than do those of few affected daughters.

To test the second component of our hypothesis, i.e., the association between maternal carrier status and daughter's breast cancer status, we evaluated whether the genotypes of mothers with more affected daughters differ from those of mothers with less affected daughters. Such deviation might be expected if some aspect of a daughter's in-utero environment, governed by the mother's genotype, influences the daughter's risk of subsequent breast cancer development. The FS was easily adapted to evaluate this question, simply by comparing separately the observed or imputed genotypes of mothers with a higher number of affected daughters to those of mothers with less number of affected daughters. A statistically significant value of the FS when restricted to the mothers (but not when restricted to the fathers) would provide evidence for this association.

When the null hypothesis is rejected, it is useful to estimate a measure of association between genotype and risk, such as the odds-ratio, and to evaluate the effects of potential confounding by hormonal factors. To do so, we also performed conditional logistic regression analyses (30, 31) on the all the available sibships containing at least one affected sib and at least

one unaffected sib who had provided blood samples and relevant epidemiology data for statistical adjustment (165 sibships for *CYP17* and 169 sibships for *CYP19*).

Results

Of the 277 nuclear families eligible for *CYP17* analyses 229 were Caucasian, 4 were African-American, 41 were Hispanic, and 3 were Asian-American. Of the 278 nuclear families eligible for CYP19 analyses 229 were Caucasian, 4 were African-American, 42 were Hispanic, and 3 were Asian-American. Table 1 shows the distribution of the study subjects according to *CYP17* and *CYP19* genotypes, by family position and breast cancer status. The numbers in each cell represent the number of specific type of family members in our study population carrying a particular genotype. The number of TTTA repeats in intron 4 of the *CYP19* gene ranged between 7 and 13 in our study population, with the 7 ($TTTA_7$) and 11 ($TTTA_{11}$) repeat alleles being the most frequent (allele frequency 53.9% and 28.8% respectively). These frequencies are consistent with those found in Caucasian populations in other studies in the U.S. (15). The frequency of the *CYP17* variant C allele was 42.8% in this study population, which is similar to other studies conducted in Caucasians (4).

The distribution of the nuclear families according to mother's and father's carrier status and mother's and daughter's affected status is presented in Table 2. The majority (~55%) of the nuclear families contained one affected and one unaffected daughter. The majority of the nuclear families had one or more parents who did not have the genotyping information available.

Table 3 presents the FS and NFS for testing the associations between the *a priori* hypothesized *CYP17* and *CYP19* variant alleles and breast cancer Each test statistic has approximately a standard gaussian distribution under the null hypothesis of no association between genotype and breast cancer risk. A positive value of a NFS reflects excess transmission of the variant allele to affected daughters, and a negative value represents fewer such transmissions than expected under the null. Thus a test statistic that is negative but large in absolute value would suggest that the variant allele is associated with reduced risk. We computed the FS and NFS under recessive, dominant and additive models. Based on the literature, we hypothesized *a priori* that *CYP19* alleles with 10 or more TTTA repeats would be associated with breast cancer. We examined the association between the *CYP19* genotype and breast cancer by defining the variant allele(s) by treating each of the ≥10 repeat alleles ($TTTA_{10}$, $TTTA_{11}$, $TTTA_{12}$ and $TTTA_{13}$) separately as the variant allele under each of the three models (realizing that this may have increased the chance of our finding of a statistically significant association - "see discussion").

As seen in Table 3, the NFS for association between $TTTA_{11}$ allele and breast cancer under the dominant model is 1.83, which is higher than the critical value (1.65) for a one-tailed test statistic, suggesting that affected daughters were more likely to receive the $TTTA_{11}$ allele from their parents (irrespective of their ethnic distribution) than unaffected daughters. Like the NFS, the FS was also statistically significant under the dominant model, supporting an association between the *CYP19* $TTTA_{11}$ allele and breast cancer among the parents in these families. None of the other *CYP19* alleles showed a consistent association with breast cancer based on the NFS and FS (results not shown). Although the FS found an association between the *CYP19* $TTTA_{13}$ allele and breast cancer, this was not supported by the more robust NFS (results not shown).

Neither the FS nor the NFS suggested any significant association between the *CYP17* variant C allele and breast cancer, under any of the models of FGAP analyses (see Table 3).

Table 4 presents the results of conditional logistic regression analysis comparing the *CYP17* and *CYP19* genotypes between affected and unaffected sisters. These results, adjusted for

age (in years), hormone replacement use (ever/never), oral contraceptive use (ever/never), age at menarche (in years) and full term pregnancies (yes/no), are similar to the FGAP results. As seen in Table 4, carriers of *CYP19* $TTTA_{11}$ allele had an increased risk of breast cancer (OR 1.8; 95% CI 0.9-3.5).

Table 5 presents results relating maternal and paternal carrier statuses for the variants of estrogen-biosynthesis genes *CYP17* and *CYP19* with breast cancer risk in daughters. Mothers of higher numbers of affected daughters were more likely to carry the *CYP19* alleles with 11 TTTA repeats ($TTTA_{11}$) than mothers of fewer unaffected daughters, although this association did not achieve statistical significance. There were no such associations between the paternal carrier status of $TTTA_{11}$ or any of the other *CYP19* alleles and breast cancer in daughters. Although maternal carrier status of the *CYP17* C allele tended to be positively associated with daughter's breast cancer this association was not statistically significant.

Discussion

Despite a sound biological basis for the role of estrogen-biosynthesis genes in breast cancer, the findings of studies investigating the relationship between these genes and breast cancer have not been consistent. Employing a case-control design, many of these prior studies, especially those examining the *CYP17* gene-breast cancer relationships, produced conflicting results. Although compared to *CYP17* a smaller number of studies investigated the association of breast cancer with *CYP19*, findings for *CYP19* have been relatively more consistent, with most studies showing a positive association between *CYP19* alleles with a higher number (10, 12 or ≥10) of TTTA repeats and breast cancer (13, 15-17).

Using a family-based design we investigated the relationships between the *CYP17* and *CYP19* gene variants and breast cancer among families participating in the MNYR. Like many of the prior case-control studies, the present study did not find any association between the *CYP17* C (variant) allele and breast cancer. However, our findings support an association between certain alleles of the *CYP19* intron 4 TTTA repeat polymorphism and breast cancer. Based on the prior studies we defined each of the *CYP19* alleles with 10, 11, 12 or 13 TTTA repeats as the 'variant' allele and examined each association with breast cancer. Unlike some of the prior case-control studies we did not find the *CYP19* $TTTA_{10}$ or $TTTA_{12}$ alleles to be associated with breast cancer. However, we found the *CYP19* $TTTA_{11}$ allele to be significantly associated with breast cancer in these nuclear families, under a dominant model. Although we also observed a significantly positive association between the *CYP19* $TTTA_{13}$ allele and breast cancer among the parents in these families, we did not observe excess transmission from parents to affected daughters, suggesting that the association may be due to chance or bias.

In addition to evaluating associations between a woman's breast cancer risk and her own constitutional genotype, we also evaluated whether maternal genotypes are associated with the breast cancer risk in the daughters (independent of daughter's own genotype). We found that the maternal (but not paternal) genotypes of *CYP19* allele with 11 TTTA repeats conferred a non-significantly elevated breast cancer risk to daughters. This association, although weak, is consistent with evidence from the prior literature on the association between in-utero exposure to hormonal factors and breast cancer risk in adulthood (18). While the association may be due to chance, if confirmed in subsequent studies, it will have important implications in advancing our understanding of the breast cancer etiology.

Some limitations of the current study warrant consideration. The major limitation concerns

statistical power. The analysis, which is based on 287 nuclear families, may not have had enough power to detect small increases in risk associated with certain of the *CYP17* genotypes. For example, we lacked power to evaluate interactions between genotypes for *CYP17* and *CYP19* and both endogenous and exogenous hormonal characteristics, such as age at menarche, timing and number of pregnancies and use of exogenous hormones. The evaluation of such interaction will be the subject of a separate future analysis, based on additional numbers of CFR families.

Although the hypotheses examined in this study are not novel, the study design (which is free from population stratification bias) and the analytical approach have not been applied to these hypotheses in previous studies. Several limitations of this study warrant caution when interpreting the findings. First, the selection of nuclear families participating in this study from the MNYR was not population-based. While this may limit the generalizability of the findings it should not affect the validity of the observed associations. Second, although it is possible for variations in the number of nucleotide repeat in hormone-related genes to be associated with cancer risk, such an association is less plausible biologically for the TTTA repeat numbers in the *CYP19* gene. This is because the TTTA polymorphism is in the intronic region of the gene and so, it is less likely that the variant alleles of the gene are directly associated with the functional status of endogenous estrogens in the body. Nevertheless, it is possible that one or more of the *CYP19* TTTA alleles, including the $TTTA_{11}$ allele, are in linkage disequilibrium with other functionally relevant alleles, as suggested by other studies (16). Third, the current study compared multiple *CYP19* TTTA alleles with breast cancer under different models. Although it is possible that multiple comparisons may have led to the observed associations, the consistency of the associations involving the *CYP19* $TTTA_{11}$ allele across both parents and transmission to offspring as well as the similarity between the associations with constitutional as well as maternal genotypes suggest that these findings may have a biological basis. Further, the fact that the association was observed under a dominant model only (not under recessive or additive models) and consistent with conditional logistic regression analysis may be suggestive of the specificity of the finding.

In conclusion, this family-based study found that the *CYP19* allele with 11 TTTA repeats ($TTTA_{11}$) is associated with breast cancer risk among families participating in a breast cancer family registry. The study also suggests that maternal carrier status of the *CYP19* $TTTA_{11}$ allele may be associated with breast cancer in daughters in these families. These associations may have important implications for understanding etiology and risk prediction of breast cancer. However, since this is the first study to report an association with the *CYP19* $TTTA_{11}$ allele, and since multiple comparisons have been made, the associations reported in this study should be interpreted with caution and need to be confirmed in future family-based studies.

Reference List

1. Dunning, A. M., Healey, C. S., Pharoah, P. D., Teare, M. D., Ponder, B. A., and Easton, D. F. A systematic review of genetic polymorphisms and breast cancer risk. Cancer Epidemiol. Biomarkers Prev., *8:* 843-854, 1999.

2. Ye, Z. and Parry, J. M. The CYP17 MspA1 polymorphism and breast cancer risk: a meta-analysis. Mutagenesis, *17:* 119-126, 2002.

3. Nedelcheva, K., V, Haraldsen, E. K., Anderson, K. B., Lonning, P. E., Erikstein, B., Karesen, R., Gabrielsen, O. S., and Borresen-Dale, A. L. CYP17 and breast cancer risk: the polymorphism in the 5' flanking area of the gene does not influence binding to Sp-1. Cancer Res., *59:* 2825-2828, 1999.

4. Feigelson, H. S., Shames, L. S., Pike, M. C., Coetzee, G. A., Stanczyk, F. Z., and Henderson, B. E. Cytochrome P450c17alpha gene (CYP17) polymorphism is associated with serum estrogen and progesterone concentrations. Cancer Res., *58:* 585-587, 1998.

5. Haiman, C. A., Hankinson, S. E., Spiegelman, D., Colditz, G. A., Willett, W. C., Speizer, F. E., Kelsey, K. T., and Hunter, D. J. The relationship between a polymorphism in CYP17 with plasma hormone levels and breast cancer. Cancer Res., *59:* 1015-1020, 1999.

6. Feigelson, H. S., Coetzee, G. A., Kolonel, L. N., Ross, R. K., and Henderson, B. E. A polymorphism in the CYP17 gene increases the risk of breast cancer. Cancer Res., *57:* 1063-1065, 1997.

7. Bergman-Jungestrom, M., Gentile, M., Lundin, A. C., and Wingren, S. Association between CYP17 gene polymorphism and risk of breast cancer in young women. Int. J. Cancer, *84:* 350-353, 1999.

8. Spurdle, A. B., Hopper, J. L., Dite, G. S., Chen, X., Cui, J., McCredie, M. R., Giles, G. G., Southey, M. C., Venter, D. J., Easton, D. F., and Chenevix-Trench, G. CYP17 promoter polymorphism and breast cancer in Australian women under age forty years. J. Natl. Cancer Inst., *92:* 1674-1681, 2000.

9. Miyoshi, Y., Iwao, K., Ikeda, N., Egawa, C., and Noguchi, S. Genetic polymorphism in CYP17 and breast cancer risk in Japanese women. Eur. J. Cancer, *36:* 2375-2379, 2000.

10. Huang, C. S., Chern, H. D., Chang, K. J., Cheng, C. W., Hsu, S. M., and Shen, C. Y. Breast cancer risk associated with genotype polymorphism of the estrogen-metabolizing genes CYP17, CYP1A1, and COMT: a multigenic study on cancer susceptibility. Cancer Res., *59:* 4870-4875, 1999.

11. Feigelson, H. S., McKean-Cowdin, R., Pike, M. C., Coetzee, G. A., Kolonel, L. N., Nomura, A. M., Le Marchand, L., and Henderson, B. E. Cytochrome P450c17alpha gene (CYP17) polymorphism predicts use of hormone replacement therapy. Cancer Res., *59:* 3908-3910, 1999.

12. Mitrunen, K., Jourenkova, N., Kataja, V., Eskelinen, M., Kosma, V. M., Benhamou, S., Vainio, H., Uusitupa, M., and Hirvonen, A. Steroid metabolism gene CYP17 polymorphism and the development of breast cancer. Cancer Epidemiol. Biomarkers Prev., *9:* 1343-1348, 2000.

13. Kristensen, V. N., Andersen, T. I., Lindblom, A., Erikstein, B., Magnus, P., and Borresen-Dale, A. L. A rare CYP19 (aromatase) variant may increase the risk of breast cancer. Pharmacogenetics, *8:* 43-48, 1998.

14. Siegelmann-Danieli, N. and Buetow, K. H. Constitutional genetic variation at the human aromatase gene (Cyp19) and breast cancer risk. Br. J. Cancer, *79:* 456-463, 1999.

15. Haiman, C. A., Hankinson, S. E., Spiegelman, D., De, V., I, Colditz, G. A., Willett, W. C., Speizer, F. E., and Hunter, D. J. A tetranucleotide repeat polymorphism in CYP19 and breast cancer risk. Int. J. Cancer, *87:* 204-210, 2000.

16. Baxter, S. W., Choong, D. Y., Eccles, D. M., and Campbell, I. G. Polymorphic variation in CYP19 and the risk of breast cancer. Carcinogenesis, *22:* 347-349, 2001.

17. Miyoshi, Y., Iwao, K., Ikeda, N., Egawa, C., and Noguchi, S. Breast cancer risk associated with polymorphism in CYP19 in Japanese women. Int. J. Cancer, *89:* 325-328, 2000.

18. Sanderson, M., Williams, M. A., Daling, J. R., Holt, V. L., Malone, K. E., Self, S. G., and Moore, D. E. Maternal factors and breast cancer risk among young women. Paediatr. Perinat. Epidemiol., *12:* 397-407, 1998.

19. Weiss, H. A., Potischman, N. A., Brinton, L. A., Brogan, D., Coates, R. J., Gammon, M. D., Malone, K. E., and Schoenberg, J. B. Prenatal and perinatal risk factors for breast cancer in young women. Epidemiology, *8:* 181-187, 1997.

20. Ekbom, A., Hsieh, C. C., Lipworth, L., Adami, H. Q., and Trichopoulos, D. Intrauterine environment and breast cancer risk in women: a population-based study. J. Natl. Cancer Inst., *89:* 71-76, 1997.

21. Potischman, N. and Troisi, R. In-utero and early life exposures in relation to risk of breast cancer. Cancer Causes Control, *10:* 561-573, 1999.

22. Ekbom, A., Trichopoulos, D., Adami, H. O., Hsieh, C. C., and Lan, S. J. Evidence of prenatal influences on breast cancer risk. Lancet, *340:* 1015-1018, 1992.

23. Sanderson, M., Williams, M. A., Malone, K. E., Stanford, J. L., Emanuel, I., White, E., and Daling, J. R. Perinatal factors and risk of breast cancer. Epidemiology, *7:* 34-37, 1996.

24. Kwok, P. Y. SNP genotyping with fluorescence polarization detection. Hum Mutat., *19:* 315-323, 2002.

25. Hsu, T. M., Chen, X., Duan, S., Miller, R. D., and Kwok, P. Y. Universal SNP genotyping assay with fluorescence polarization detection. Biotechniques, *31:* 560, 562, 564-560,8, passim, 2001.

26. Whittemore, A. S. and Tu, I. P. Detection of disease genes by use of family data. I. Likelihood-based theory. Am. J. Hum Genet., *66:* 1328-1340, 2000.

27. Spielman, R. S., McGinnis, R. E., and Ewens, W. J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am. J. Hum Genet., *52:* 506-516, 1993.

28. Schaid, D. J. General score tests for associations of genetic markers with disease using cases and their parents. Genet. Epidemiol., *13:* 423-449, 1996.

29. Spielman, R. S. and Ewens, W. J. The TDT and other family-based tests for linkage disequilibrium and association. Am. J. Hum Genet., *59:* 983-989, 1996.

30. Hosmer, D. and Lemeshow, S. Logistic regression for matched case-control studies. Applied logistic regression, pp. 187-213. New York: John Wiley & Sons, 1989.

31. Breslow, N. E. and Day, N. E. Statistical methods in cancer research. Volume I - The analysis of case-control studies. IARC Sci. Publ., 5-338, 1980.

Table 1. Genotype Distribution of the Study Population by Gender, Family Relationship and Affected Status

	Mothers		Daughters		Fathers	Sons	Total
	Affected	Unaffected	Affected	Unaffected			
CYP17							
CC	2	10	47	41	6	11	117
CT	8	32	147	108	30	22	347
TT	2	19	107	61	8	18	215
Unknown	64	140	0	0	233	0	437
Total	76	201	301	210	277	51	1116
***CYP19* (# of TTTA repeats)**							
7/7	1	15	88	55	7	21	187
7/8	0	10	34	29	4	5	82
7/9	0	0	1	0	1	0	2
7/10	0	1	5	6	2	0	14
7/11	9	19	106	71	14	19	238
7/12	0	2	9	8	0	0	19
7/13	0	0	0	0	1	0	1
8/8	0	0	4	1	0	0	5
8/10	0	1	3	2	0	0	6
8/11	1	2	24	16	1	3	47
8/12	0	0	1	2	0	0	3
8/13	0	0	0	1	0	0	1
9/11	0	0	0	1	0	0	1
10/11	0	0	2	3	0	2	7
11/11	0	5	19	16	1	1	42
11/12	0	1	4	2	0	0	7
11/13	1	0	2	1	1	0	5
11/not 11*	1	3	0	0	5	0	9
Unknown	63	143	0	0	241	0	447
Total	76	202	302	214	278	51	1123

*Indicates those whose genotype cannot be inferred for both alleles; the other allele could be 7, 8, or 12. Two of these 9 observations, one is a unaffected mother and the other is the father in the same nuclear family, will be excluded when allele with ≥10 repeats is selected as bad allele, because either them could be 11/12.

Table 2. Distribution of Participating Nuclear Families According to Mother's Breast Cancer Status, Mother's and Father's Carrier Status of the *CYP17* and *CYP19* Variant Alleles, and Number of Affected/Unaffected Daughters

Number of Daughters	Number of Nuclear Families According to Mother's Breast Cancer Status and Genotype						Total	Number of Nuclear Families According to Father's Genotype			Total
	Affected			Unaffected							
Affected/Unaffected	Carrier	Non-carrier	Unknown	Carrier	Non-carrier	Unknown		Carrier	Non-carrier	Unknown	
CYP17											
1/0	3	0	19	17	12	29	80	12	2	66	80
1/1	5	2	38	16	6	85	152	19	4	129	152
1/2	0	0	2	4	0	14	20	4	2	14	20
1/3	0	0	1	2	0	0	3	0	0	3	3
2/0	2	0	3	1	1	9	16	0	0	16	16
2/1	0	0	0	0	0	2	2	0	0	2	2
2+/2+	0	0	1	2	0	1	4	1	0	3	4
Total	10	2	64	42	19	140	277	36	8	233	277
CYP19											
1/0	4	0	17	14	13	29	77	7	4	66	77
1/1	5	1	40	12	9	88	155	10	6	139	155
1/2	1	0	1	5	1	14	22	5	2	15	22
1/3	1	0	0	1	0	0	2	1	0	1	2
2/0	1	0	4	1	1	9	16	1	0	15	16
2/1	0	0	0	0	0	2	2	0	0	2	2
2+/2+	0	0	1	0	1	2	4	0	0	4	4
Total	12	1	63	33	25	144	278	24	12	242	278

Table 3. Association Between the *CYP19* and *CYP17* Variant Alleles and Breast Cancer

Variant Allele(s)	Estimated Allele Frequency (%)	Nonfounder Statistic[2]			Founder Statistic[2]		
		Recessive Model	Dominant Model	Additive Model	Recessive Model	Dominant Model	Additive Model
CYP17							
C	42.46	-1.01	-1.52	-1.85	0.40	1.08	1.01
(p-value)[1]		(0.16)	(0.06)	(0.03)	(0.34)	(0.14)	(0.16)
CYP19 (# of TTTA repeats)							
$(TTTA)_{11}$	28.78	-1.09	**1.83**	0.97	-1.50	**1.96**	0.89
(p-value)[1]		(0.14)	(0.03)	(0.17)	(0.07)	(0.03)	(0.19)

[1] Based on one-tailed test statistics

[2] Values that are statistically significant at one-tailed test are displayed in 'bold'

Table 4. Conditional Logistic Regression Analysis of Discordant Sib-ships for the Association between *CYP17* and *CYP19* Genotypes and Breast Cancer

Gene (sibling sets/cases/controls)[2]		Affected (N)	Unaffecte (N)	Adjusted Odds Ratios for Beast Cancer[1]
CYP17 **(165/171/188)**				
Recessive model	TC/TT	146	154	1.0
	CC	25	34	0.6 (0.3-1.4)
Additive model	TT	59	56	1.0
	TC	87	98	0.9 (0.5-1.6)
	CC	25	34	0.6 (0.2-1.4)
Dominant model	TT	59	56	1.0
	TC/CC	112	132	0.8 (0.5-1.5)
CYP19 (# of TTTA repeats)**(169/175/193)**				
Recessive model	$(TTTA)_{11}(TTTA)_{other}/(TTTA)_{11}(TTTA)_{other}$	165	179	1.0
	$(TTTA)_{11}(TTTA)_{11}$	10	14	0.7 (0.2-2.3)
Additive model	$(TTTA)_{other}(TTTA)_{other}$	77	95	1.0
	$(TTTA)_{11}$ $(TTTA)_{other}$	88	84	1.8 (0.9-3.6)
	$(TTTA)_{11}$ $(TTTA)_{11}$	10	14	1.1 (0.3-4.1)
Dominant model	$(TTTA)_{other}(TTTA)_{other}$	77	95	1.0
	$(TTTA)_{11}(TTTA)_{other}/(TTTA)_{11}(TTTA)_{11}$	98	98	1.8 (0.9-3.5)

[1] Odds ratios were adjusted for age (in years), hormone replacement use (ever/never), oral contraceptive use (ever/never), age at menarche (in years), full term pregnancies (yes/no)

[2] Each sibling set had at least one breast cancer case and one sister control. All the subjects included in the analysis had information for all the covariate variables.

* $p < 0.05$

Table 5. Association Between Parental Carrier Status of the Variant Allele(s) and Breast Cancer Risk in Daughters

Variant Allele(s)	Test Statistic[1]	
	Mothers' Carrier Status and Disease Risk in Daughters	Fathers' Carrier Status and Disease Risk in Daughters
CYP17		
C	1.47	1.40
(p-value)[2]	(0.07)	(0.08)
CYP19 (#of repeats)		
$(TTTA)_{11}$	1.52	-0.11
(p-value)[2]	(0.06)	(0.46)

[1]Calculated under additive model

[2]Based on one-tailed test statistics

Appendix 2

Breast Cancer Research and Treatment **85:** 121–131, 2004.

Report

A family-based genetic association study of variants in estrogen-metabolism genes *COMT* and *CYP1B1* and breast cancer risk

Habibul Ahsan[1,3], Yu Chen[1], Alice S. Whittemore[4], Muhammad G. Kibriya[1], Irina Gurvich[2], Ruby T. Senie[1,3], and Regina M. Santella[2,3]

[1]*Department of Epidemiology,* [2]*Department of Environmental Health Sciences, Mailman School of Public Health,* [3]*Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY;* [4]*Department of Health Research and Policy, Division of Epidemiology, Stanford University, Stanford, CA, USA*

Key words: breast cancer, *COMT*, *CYP1B1*, estrogen-metabolism genes, genetic epidemiology, genetic polymorphism

Summary

In this paper, we report findings from a family-based association study examining the association between polymorphisms in two key estrogen-metabolism genes *CYP1B1* (codon 432 G→C and codon 453 A→G variants) and *COMT* (codon 158 G→A variant) and female breast cancer. We conducted the study among 280 nuclear families containing one or more daughters with breast cancer with a total of 1124 family members (702 with available constitutional DNA and questionnaire data and 421 without). These nuclear families were selected from breast cancer families participating in the Metropolitan New York Registry (MNYR) – one of the six centers of NCI's Breast Cooperative Family Registry. We used likelihood-based statistical methods to examine the allelic associations. We found none of the variant alleles of the *CYP1B1* and *COMT* genes to be associated with breast cancer in these families. This was consistent with results from matched case-control analyses using all available sib-ships in these families. However, we found that parental carrier status of the *CYP1B1* codon 453 variant G allele and the *COMT* codon 158 variant A allele was associated with breast cancer risk in daughters (independent of the daughters' own genotype). In conclusion, findings from this family-based study indicate that a woman's own *CYP1B1* or *COMT* genotypes are not associated with her breast cancer risk. Although the study found that parental carrier status of certain *CYP1B1* or *COMT* genotypes might be associated with daughter's breast cancer risk, the biological basis as well as independent confirmation of this finding need to be investigated in future larger family-based studies before making meaningful inferences.

Introduction and background

Cumulative exposure to circulating estrogen is considered to be of primary importance in breast cancer etiology. Estrogen biosynthesis, cellular binding and metabolism involve many steps and the genes controlling these steps may contribute to inherent variability in breast cancer susceptibility. Endogenous estrogen is produced predominantly in the ovarian theca cells in premenopausal women and in the breast stromal adipose cells in postmenopausal women but metabolized in different tissues. The current study focuses on *CYP1B1* and *COMT*, two key genes that control the metabolism of estradiol and estrones once they are synthesized from their lipid precursors.

CYP1B1 catalyzes the conversion of estradiol to 4-OH catechol estrogen metabolites. The 4-OH catechol estrogen metabolites have been postulated to be capable of initiating mammary tumors through their reactive metabolites [1]. Recently, reactive metabolites of 4-OH catechol estrogens, that is, 3- and 4-OH quinones and semiquinones have been shown to initiate mammary tumors in animals through formation of depurinating DNA adducts which are capable of

creating *de novo* mutations [2]. CYP1B1 expression is altered by the 2,3,7,8-tetrachloro dibenzo-*p*-dioxin (TCDD) [3] and has been immunohistochemically detected in breast and other tumor tissues [4–6] but not in normal tissues [7]. The gene encoding human CYP1B1 is located in the chromosome 2p21–22 region and contains three exons and two introns [3, 8].

Two *CYP1B1* polymorphisms, m1 and m2, resulting from base substitutions in codon 432 (Val → Leu) and 453 (Asn→Ser) respectively, have been detected in exon 3 which encodes the catalytically important heme-binding domain of the enzyme. The functional relevance of these variants in relation to disease risk is not clearly understood. To date, four published studies have examined the associations between *CYP1B1* variants and breast cancer. Bailey et al. first examined the hypothesis among 219 Caucasian and African-American cases and their frequency-matched (on age and race) controls [9]. The frequencies of m1 and m2 alleles differed significantly between the Caucasian and African-American subjects in this study. Approximately 40% of the Caucasian women carried the M1 Val allele as compared to 70% of the African-American women. Similarly, 17.4% of the Caucasian women carried the rare m2 Serine allele as compared to only 3.4% of the African-Americans. Although the authors did not find a significant association between these *CYP1B1* codon 432 and codon 453 variants and breast cancer, the codon 432 variant C allele was significantly associated with the estrogen and progesterone receptor negative breast cancers [9]. Zheng et al., in a case-control study in China found an increased risk of breast cancer in relation to the codon 432 C allele [10]. Similar associations between the codon 432 variant C allele and breast cancer were also found by Watanabe et al. among Japanese women [11] and by Kocabas and colleagues among Turkish women [12].

COMT detoxifies the catechol estrogen metabolites (4-OH estradiols/estrones) which have been postulated to be capable of initiating mammary tumors through their reactive metabolites by formation of depurinating DNA adducts which are capable of creating *de novo* oncogenic mutations [2]. The COMT enzyme catalyzes the transfer of a methyl group from the *S*-adenosyl-L-methionine to the *m*-hydroxy group of catechol compounds, rendering the catechol estrogens more water soluble and enhancing excretion from the body [1]. The gene that encodes COMT is polymorphic, with a G to A transition in exon 4, resulting in a valine to methionine substitution at codon 158 [13]. About 25% of women in the general population are homozygous for this variant allele. The variant allele is heat labile and has a six-fold lower activity than the normal wild-type allele. Women with this low activity allele thus have a lower excretion of carcinogenic catechol estrogen metabolites and thus may be at a higher risk of breast cancer. A nested case-control study in Maryland first showed that postmenopausal women homozygous for the *COMT* low-activity allele (AA genotype) are more than twice as likely to get breast cancer than those carrying at least one normal allele (GG or GA genotype) [14]. After publication of this finding, several other case-control studies have also examined this association. The Western New York study found a two-fold (OR 2.1, 95% CI 1.4–4.3) increased risk among premenopausal women [15]. The Carolina Breast Cancer Study however, found no substantial increased risk in relation to the low-activity genotype [16]. The authors of the original article showed a higher risk in relation to *COMT* low-activity allele among Maryland women with low levels of serum folate [17]. Three of the four published studies among Asian women found the variant *COMT* low-activity allele to be significantly associated with breast cancer (OR 4.0, 95% CI 1.1–19.1 among Taiwanese [18]; OR 1.7, 95% CI 1.0–2.8 among Korean [19]; and OR 4.0, 95% CI 1.1–19.1 [20] and OR 1.0, 95% CI 0.5–2.0 [21] among Japanese). On the other hand, two published studies among Scandinavian women found an inverse association (OR 0.4, 95% CI 0.2–0.9 among Finish women and OR 0.8, 95% CI 0.4–2.0 among Swedish women) [22, 23].

All published studies examining the association between the *CYP1B1* and *COMT* polymorphisms and breast cancer used a classical case-control design, which may be susceptible to population stratification bias. In this study, using a family-based design, we evaluated the hypotheses that breast cancer risk is elevated among carriers of the *CYP1B1* codon 432 G → C and codon 453 A → G variants, which are associated with increased activity of CYP1B1, and carriers of *COMT* codon 158 G → A variant, which is associated with decreased COMT activity among nuclear families participating in the Metropolitan New York Registry (MNYR) – one of the six international centers of U.S. NCI's Breast Cancer Family Registry (BCFR) project. All published studies focused on the relationship between a woman's own constitutional genotype and her breast cancer risk. A body of recent literature has provided limited data suggesting that a woman's breast cancer risk may be related not only

to her own endogenous estrogens during adolescence and adulthood, but also to her prenatal exposure, that is, *in utero* exposure to her maternal circulating estrogens [24–29]. In addition to the main hypotheses about the association between a woman's own genotype and her breast cancer status, the family-based design of the current study allowed us to address this additional hypothesis indirectly, by examining the association between maternal carrier status of *CYP1B1* or *COMT* gene variants (i.e., *in utero* exposure to an altered level of maternal estrogens) and breast cancer status in daughters.

Methods

The descriptions of the sources of study participants, their recruitment and data collection methods as part of the parent MNYR project have been described in details elsewhere [http://www.ccc.columbia.edu/metro.html] [30]. Briefly, the MNYR has been recruiting high risk breast and/or ovarian cancer families from clinical and community settings within the metropolitan New York area since 1995 who met the following criteria: families having one or more members with breast cancer or ovarian cancer diagnosed at <45 years of age; one or more members with both breast and ovarian cancer; two or more relatives with breast or ovarian cancer diagnosed at age ≥45 years, or any male with breast cancer. This unique resource can be used to identify new avenues for prevention, detection and treatment of breast cancer. After identification of a potentially eligible family through a proband, he/she is invited to participate in the registry and his/her family's eligibility is assessed. If the proband agrees to participate and his/her family meets the eligibility criteria, with appropriate informed consent, he/she is interviewed either in-person or by phone with epidemiology and family-history questionnaires. A self-administered dietary questionnaire is also provided for return by mail. A brief follow-up questionnaire will be completed each year. In addition, a sample of blood or buccal cells is also collected from the participants. Once the proband provides permission to contact family members, they are invited to participate. If the family members consent to participate, questionnaire data, blood or buccal samples from them are also collected in a similar manner. Tumor tissue samples are collected and reviewed histopathologically for participants affected with cancer. Participants in MNYR are provided information newsletters and are invited to attend educational seminars. Participants have access to referrals to the Cancer Genetic Program(s) in and around New York City for counseling and genetic testing are provided upon request. To date, the MNYR has enrolled nearly 1200 families and nearly 4000 total participants.

For the present family-based study, we included nuclear families having at least one affected daughter with DNA and at least one parent and/or sibling with DNA. Altogether, 280 families met these eligibility criteria. Respondents can participate in the MNYR with or without completion of the full epidemiology questionnaire and/or blood samples. There were 1123 family members in the 280 eligible nuclear families of which 702 completed the full epidemiology questionnaire and provided blood samples. However, accurate data on relevant variables for the statistical method used in this study (see below) for the remaining 421 members were available from the family-history questionnaire completed by the 702 members. There was 99% concordance in data on age and affected status between women who completed the full epidemiology questionnaire and women who did not.

In this study, we typed the participants for the *CYP1B1* codon 432 G→C (NCBI rs# 1056836) and codon 453 A→G variants (NCBI rs# 1800440) and the *COMT* codon 158 G→A variant (NCBI rs# 4680). A total of 26 subjects could not be typed for the *CYP1B1* variants and 23 subjects could not be genotyped for the *COMT* variant. Thus, genotype data were available on a total of 676 members (from 278 nuclear families) for *CYP1B1* and 679 members (from 277 nuclear families) for *COMT*.

The genotyping was performed using template-directed primer extension with detection of incorporated nucleotides by fluorescence polarization in a 96-microwell-based format, essentially as described [31, 32]. Master DNA 96-well plates containing 10 μg/ml were used to make replica plates containing 25 ng DNA/well. First, genomic DNA was amplified by polymerase chain reaction (PCR) using the appropriate primers as follows: forward 5′- TGT CAA CCA GTG GTC TGT GAA -3′ and reverse 5′- AAA TCA TCA CTC TGC TGG TCA -3′ for the *CYP1B1* codon 432 G→C SNP; forward 5′- GAC CCA GTG AAG TGG CCT AA -3′ and reverse 5′- GCC AGG ATG GAG ATG AAG AG -3′ for the *CYP1B1* codon 453 A→G SNP; and forward 5′- TCA CCA TCG AGA TCA ACC CC -3′ and reverse 5′- ACA ACG GGT CAG GCA TGC A -3′ for the *COMT* codon 158 G→A SNP.

PCR was carried out in 10 μl reaction mixture containing 25 ng of genomic DNA, 1.6 pmols of each forward and reverse primers, 1 μl of 10× PCR buffer, 0.5 Units of Taq polymerase (Roche Molecular Biochemicals, Indianapolis, IN), 0.25 μl of 10 mM dNTPs (Roche) and 5.75 μl water. The thermocycling conditions for all three SNPs were: initial denaturation for 6 min at 94°C, followed by 35 cycles of denaturation for 30 s at 94°C, annealing for 45 s at specific temperature (55°C for *CYP1B1* codon 432 G→C and *CYP1B1* codon 453 A→G and 60°C for *COMT* codon 432 A→G SNPs) and 1 min extension at 72°C, with a final extension at 72°C 4 min. After PCR amplification (125 bp for *CYP1B1* codon 432 G→C, 176 bp for *CYP1B1* codon 453 A→G, and 96 bp amplicon for *COMT* codon 158 G→A), the primers and dNTPs were digested with the Shrimp alkaline cocktail containing 1 unit of shrimp alkaline phosphatase (Roche), 1 μl of 10× buffer and 1 unit of *E. Coli* exonuclease I (10 u/μl, United States Biochemical, Cleveland, OH) and 7.9 μl of water for 45 min at 37° followed by heating at 95° for 15 min to deactivate the enzyme. Then, single nucleotide extension was carried out in the presence of the appropriate allele specific ddNTPs differentially fluorescence-labeled with either R110 or TAMRA (Acycloprime FP SNP Detection kits G/C, C/T and G/A for the *CYP1B1* codon 432 G→C, *CYP1B1* codon 453 A→G, and *COMT* codon 158 G→A SNPs, respectively, purchased from Perkin Elmer Life Sciences, Boston, MA). For the single nucleotide extension reaction, both forward and reverse probes were tested to select the optimum based on clear signal differences. The optimum chosen were the reverse extension primer 5′- AAA GTT CTC CGG GTT AGG CCA CTT CA -3′ for *CYP1B1* codon 432 G→C, the reverse extension primer 5′- CTC TGC TGG TCA GGT CCT TG -3′ for *CYP1B1* codon 453 A→G and the forward extension primer 5′- GGA TGG TGG ATT TCG CTG GC -3′ for *COMT* codon 158 G→A SNPs. Reaction mixture (13 μl/well) containing 0.05 μl Acycloprime enzyme, 1 μl Terminatior mix, 2 μl 10× reaction buffer, 0.5 μl extension primer (10 pmol/μl) and 9.45 μl water was added to 7 μl of PCR product. Single base extension was carried out under specific conditions for the *CYP1B1* codon 432 G→C and *CYP1B1* codon 453 A→G (heating at 95°C for 2 min followed by 30 cycles of 95°C for 15 s and 55°C for 30 s) and *COMT* codon 158 G→A (heating at 95°C for 2 min followed by 25 cycles of 95°C for 15 s and 55°C for 30 s) SNPs. Finally, the fluorescence polarization was measured with a Perkin Elmer Victor instrument. All the assays were performed with laboratory personnel blinded to the subject's disease status or family relationships. In addition to assay specific quality control samples, 10% of samples were re-assayed after re-labeling to keep laboratory staff blinded to its identity.

We employed the Family Genetic Analysis Program (FGAP [33, 34]) to test the null hypothesis of no association between genotype and breast cancer risk in nuclear families. The FGAP computes two test statistics: (1) the *nonfounder statistic* (NFS), a generalization of the transmission disequilibrium test (TDT) [35, 36] which evaluates transmission disequilibrium from parents to offspring, and (2) the *founder statistic* (FS), which compares the distribution of parental genotypes to that expected under the null hypothesis of no association. The FGAP statistics fully exploit data from families with variable numbers of affected/unaffected members with variable (known/unknown) patterns of parental genotypes.

Based on the prior evidence [9–12, 14–16, 18, 19, 21–23] we hypothesized that breast cancer risk is elevated among carriers of the variant alleles in two key estrogen-metabolism genes (*CYP1B1* codon 432 G→C and codon 453 A→G variants and *COMT* codon 158 G→A variant). The data analysis was focused on two specific components of the study hypotheses: (1) whether a woman's carrier status of the hypothesized alleles is associated with her breast cancer status, (2) whether parental carrier status of the hypothesized alleles is associated with their daughter's breast cancer risk.

For testing the first component of our hypotheses, we applied the FS and NFS to assess whether specific genotypes of each of the studied genes are related to breast cancer. Since FS and NFS follow a normal Gaussian distribution under the null hypothesis, the assessment of statistical significance of the association can be done based on the deviation of these statistics from the standard critical values under normal distribution. The detailed descriptions for the calculations of NFS and FS and their utility and interpretations are provided elsewhere [30] and will only be briefly mentioned here. For simplicity, we describe these analyses for the *COMT* codon 158 G→A variant as applied to nuclear families consisting of two parents and at least one daughter. Parents may be untyped and the mother's breast cancer status may be unknown. The test statistics, which are likelihood-based score statistics, are obtained by summing the score

Table 1. Genotype distribution of the study population by gender, family relationship and affected status

	Mothers		Daughters		Fathers	Sons	Total
	Affected	Unaffected	Affected	Unaffected			
CYP1B1 codon 432							
GG	3	18	106	63	10	19	219
GC	8	25	137	104	26	23	323
CC	2	14	58	40	3	8	125
Unknown	62	145	0	0	238	0	445
Total	75	202	301	207	277	50	1112
CYP1B1 codon 453							
AA	7	41	214	157	19	37	475
AG	4	11	80	48	12	11	166
GG	0	3	10	4	3	3	23
Unknown	66	148	0	0	246	0	460
Total	77	203	304	209	280	51	1124
COMT							
GG	3	9	81	49	4	11	157
GA	8	43	148	101	27	23	350
AA	2	8	71	52	11	16	160
Unknown	63	141	0	0	235	0	439
Total	76	201	300	202	277	50	1106

contributions from each family in three steps. In the first step, we impute a probability distribution for the genotypes of each pair of parents with unknown genotype based on maximum likelihood estimates of the genotypes (GG, GA and AA), conditional on the observed genotypes of all family members. These estimates do not require the assumption of Hardy–Weinberg frequencies for parental genotypes. In the second step, we used the parental genotype distribution (known or inferred) and the offspring's observed genotypes to evaluate disequilibrium in the transmission of *COMT* A from parents to daughters. This evaluation was done via the NFS, which under the null hypothesis of no transmission disequilibrium from parent to offspring has an asymptotic standard Gaussian distribution. The NFS generalizes the TDT to families with untyped parents and to families with both affected and unaffected daughters. It is more powerful than the sibling TDT (S-TDT) test [37] when applied to families without unaffected daughters. In the final step, we used the inferred parental genotypes (and the mothers' breast cancer phenotypes) in the FS to compare the parental genotype distribution to the expected distribution under the null hypothesis of no association. This statistic treats the affected and unaffected mothers like cases and controls in a case-control study. However, each parent's contribution is weighted in proportion to his/her number of affected and unaffected daughters, so that parents of many affected daughters receive higher weights than do those of few affected daughters.

To test the second component of our hypothesis, that is, the association between maternal carrier status and daughter's breast cancer status, we evaluated whether the genotypes of mothers with more affected daughters differ from those of mothers with less affected daughters. Such deviation might be expected if some aspect of a daughter's *in utero* environment, governed by the mother's genotype, influences the daughter's risk of subsequent breast cancer development. The FS was easily adapted to evaluate this question, by comparing separately the observed or imputed genotypes of mothers with a higher number of affected daughters to those of mothers with less number of affected daughters. We also examined the association between father's carrier status and daughter's breast cancer risk. However, we did not have any *a priori* hypothesis to find an association.

In addition to testing the null hypothesis, it is useful to estimate a measure of association between

Table 2. Distribution of participating nuclear families of the *CYP1B1* codon 432/codon 453 and *COMT* variant alleles, and number of affected/unaffected daughters according to mother's breast cancer status, mother's and father's carrier status

Number of daughters	Number of nuclear families according to mother's breast cancer status and genotype						Total	Number of nuclear families according to father's genotype			Total
	Affected			Unaffected				Carrier	Noncarrier	Unknown	
Affected/ Unaffected	Carrier	Noncarrier	Unknown	Carrier	Noncarrier	Unknown					
CYP1B1 codon 432											
1/0	2	1	18	16	13	32	82	8	4	70	82
1/1	6	1	38	18	2	86	151	12	5	134	151
1/2	1	0	1	1	2	14	19	5	1	13	19
1/3	0	0	1	1	1	0	3	1	0	2	3
2/0	1	1	3	2	0	9	16	2	0	14	16
2/1	0	0	0	0	0	2	2	0	0	2	2
2+/2+	0	0	1	1	0	2	4	1	0	3	4
Total	10	3	62	39	18	145	277	29	10	238	277
CYP1B1 codon 453											
1/0	2	1	20	9	20	31	83	6	8	69	83
1/1	1	4	40	5	14	89	153	7	8	138	153
1/2	1	0	1	0	2	15	19	1	3	15	19
1/3	0	0	1	0	2	0	3	1	0	2	3
2/0	0	2	3	0	2	9	16	0	0	16	16
2/1	0	0	0	0	0	2	2	0	0	2	2
2+/2+	0	0	1	0	1	2	4	0	0	4	4
Total	4	7	66	14	41	148	280	15	19	246	280
COMT											
1/0	2	2	19	24	5	34	86	14	0	72	86
1/1	5	1	38	19	4	82	149	19	3	127	149
1/2	1	0	1	3	0	14	19	3	1	15	19
1/3	0	0	1	1	0	0	2	0	0	2	2
2/0	2	0	3	2	0	8	15	1	0	14	15
2/1	0	0	0	1	0	1	2	1	0	1	2
2+/2+	0	0	1	1	0	2	4	0	0	4	4
Total	10	3	63	51	9	141	277	38	4	235	277

genotype and risk, such as the odds-ratio, and to evaluate the effects of potential confounding by hormonal factors. To do so, we also performed conditional logistic regression analyses [38, 39] on all the available sib-ships containing at least one affected sib and at least one unaffected sib who had provided blood samples and relevant epidemiology data for statistical adjustment (165 sib-ships for *CYP1B1* codon 453 A → G, 163 sib-ships for *CYP1B1* codon 432 G → C, and 160 for *COMT*).

Results

Of the 280 nuclear families eligible for *CYP1B1* codon 453 analyses 232 were Caucasian, 4 were African-American, 41 were Hispanic, and 3 were Asian-American. Of the 277 nuclear families eligible for the *CYP1B1* codon 432 and *COMT* analyses 229 were Caucasian, 4 were African-American, 41 were Hispanic, and 3 were Asian-American. Table 1 shows the distribution of the study subjects according to the

Table 3. Association between the *CYP1B1* codon 432/codon 453 and *COMT* variant alleles and breast cancer

Allele type	Association between CYP17/CYP19 variant alleles and breast cancer						
	Estimated population allele frequency (%)	Nonfounder Statistics			Founder Statistics		
		Recessive model	Dominant model	Additive model	Recessive model	Dominant model	Additive model
CYP1B1 codon 432							
C	42.77	0.32	−0.83	−0.43	1.35	0.44	1.12
(*p*-value)		(0.38)	(0.20)	(0.33)	(0.09)	(0.33)	(0.13)
CYP1B1 codon 453							
G	16.09	0.59	0.43	0.64	*2.33*	1.09	*1.77*
(*p*-value)		(0.28)	(0.33)	(0.26)	(0.01)	(0.14)	(0.04)
COMT							
A	49.40	−0.48	−1.01	−1.08	1.47	1.30	*1.81*
(*p*-value)		(0.32)	(0.16)	(0.14)	(0.07)	(0.10)	(0.04)

CYP1B1 and *COMT* genotypes, by family position and breast cancer status. Of those with known genotypes, the frequencies of the *CYP1B1* codon 432 variant C and codon 453 variant G alleles were 43.0% and 16.0%, respectively, in our study population. The frequencies of these variant alleles appeared to be similar across relative types and affected status. Similarly, of those with known genotypes, the frequency of the *COMT* A allele was 50.2% in our study population with similar distribution across relative types and affected status. These frequencies were consistent with those found in Caucasian populations in other studies in the U.S. [9, 14–16].

Table 2 shows the distribution of the nuclear families according to mother's and father's carrier status and mother's and daughter's affected status. The majority (~54%) of the nuclear families contained one affected and one unaffected daughter. The majority of the nuclear families had one or more parents who did not have genotyping information available. These patterns were similar for distributions for all three variants.

Table 3 presents the FS and NFS for testing the associations between the *a priori* hypothesized *CYP1B1* codon 432 variant C and codon 453 variant G alleles and *COMT* codon 158 variant A allele and breast cancer. Each test statistic has approximately a standard Gaussian distribution under the null hypothesis of no association between genotype and breast cancer risk. A positive value of a NFS reflects excess transmission of the variant allele to affected daughters, and a negative value represents fewer such transmissions than expected under the null. Thus a test statistic that is positive and large would suggest that the variant allele is associated with increased risk; similarly, a test statistic that is negative but large in absolute value would suggest that the variant allele is associated with reduced risk. We computed the FS and NFS under recessive, dominant and additive models (realizing that this may have increased the chance of our finding of a statistically significant association – 'see discussion section').

As seen in Table 3, based on the NFS, none of the three variant alleles was associated with breast cancer under any of the risk models. The FSs for the association between the *CYP1B1* codon 453 variant C allele and breast cancer were statistically significant (2.33 and 1.77 under the recessive and additive models, respectively, which are higher than the critical value 1.65) for a one-tailed test statistic. Similarly, the FS for the association between the *COMT* codon 158 variant A allele and breast cancer was statistically significant (1.81 under the additive model, which is higher than the critical value 1.65) for a one-tailed test statistic. However, in the absence of any association based on NFS (which is more robust) the associations based on FS are unlikely to be meaningful.

Table 4 presents the results of conditional logistic regression analysis comparing the *CYP1B1* Codon 432/Codon 453 and *COMT* genotypes between affected and unaffected sisters. Trends in Odds ratios for *CYP1B1* codon 432 analyses seem to be more in accord with results of FS. These results, adjusted for age (in years), hormone replacement use (ever/never), oral

Table 4. Conditional logistic regression analysis of discordant sib-ships for the association between *CYP1B1* codon 432/codon 453 and *COMT* genotypes and breast cancer

Gene (number of sibling sets/ cases/controls)[a]	Affected	Unaffected	Adjusted odds ratios for breast cancer[b]
CYP1B1 codon 453 (*165/171/187*)			
Recessive model			
AA/AG	167	183	1.0
GG	4	4	1.4 (0.3–7.0)
Additive model			
AA	119	140	1.0
AG	48	43	1.9 (0.9–4.0)
GG	4	4	2.5 (0.4–14.3)
Dominant model			
AA	119	140	1.0
AG/GG	52	47	1.9 (0.9–4.1)
CYP1B1 codon 432 (*163/169/185*)			
Recessive model			
GG/GC	136	148	1.0
CC	33	37	1.0 (0.5–2.0)
Additive model			
GG	60	57	1.0
GC	76	91	0.8 (0.4–1.5)
CC	33	37	0.8 (0.3–2.0)
Dominant model			
GG	60	57	1.0
GC/CC	109	128	0.8 (0.4–1.5)
COMT (*160/166/181*)			
Recessive model			
GG/GA	124	136	1.0
AA	42	45	1.1 (0.6–2.1)
Additive model			
GG	46	44	1.0
GA	78	92	0.7 (0.3–1.3)
AA	42	45	0.8 (0.3–1.9)
Dominant model			
GG	46	44	1.0
GA/AA	120	137	0.7 (0.3–1.4)

[a] Each sibling set had at least one breast cancer case and one sister control. All the subjects included in the analysis had information for all the covariate variables.

[b] Odds ratios were adjusted for age (in years), hormone replacement use (ever/never), oral contraceptive use (ever/never), age at menarche (in years), full term pregnancies (yes/no).

contraceptive use (ever/never), age at menarche (in years) and full term pregnancies (yes/no), are consistent to the FGAP results that none of the three variant alleles was statistically significantly associated with breast cancer under any of the risk models. Table 5 presents results relating maternal and paternal carrier statuses for the variants of estrogen-metabolism genes *CYP1B1* and *COMT* with breast cancer risk in daughters. Mothers (and fathers) of higher numbers of affected daughters were more likely to carry the *CYP1B1* codon 453 G allele and the *COMT* codon 158 variant A allele than mothers (and fathers) of fewer unaffected daughters. Although parental carrier status of the *CYP1B1* codon 432 C allele tended to be positively associated with daughter's breast cancer this association was not statistically significant.

Table 5. Association between mothers' and fathers' carrier status of the variant allele(s) and breast cancer risk in daughters[a]

Variant allele(s)	Mothers' carrier status	Fathers' carrier status
CYP1B1 codon 432		
C	1.47	1.58
(*p*-value)	(0.07)	(0.06)
CYP1B1 codon 453		
G	1.95	3.17
(*p*-value)	(0.03)	(<0.001)
COMT		
A	1.90	2.45
(*p*-value)	(0.03)	(0.01)

[a] Based on FS calculated under additive model.

Discussion

Given the established hormonal roles in breast cancer etiology and the demonstrated ability for the estrogen metabolites to initiate mammary carcinogenesis, the hypothesis for the relationship between variants in genes controlling estrogen metabolism and breast cancer risk has a strong biologically basis. However, in spite of the strong biological basis, the findings from published case-control studies examining this hypothesis have been inconsistent. In this study, we investigated the hypothesis by employing a family-based design.

Our study did not find any association between any of the three variant alleles in the *CYP1B1* and *COMT* genes and breast cancer. The *CYP1B1* codon 453 G allele and the *COMT* codon 158 variant A allele were associated with breast cancer based on the FS (which is susceptible to population stratification bias) under certain genetic models. However, given the absence of

an association based on statistically more robust NFS (against population stratification bias) under any of the studied genetic models, we do not believe that these associations based on FS are real.

Three of the five prior published studies, all in non-US populations, found a positive association between the *CYP1B1* codon 432 C allele and breast cancer [9–12, 40]. The reasons for the discrepancy between our study and prior studies are not clear. If the lack of association observed in our study is real then one possible explanation would be that potential biases (especially population stratification bias) may have produced a spurious association in case-control studies to which our methodologically robust family-based design is less susceptible. However, the fact that the three positive studies were done in three different populations (Chinese, Japanese and Turkish) and two negative studies were done in the U.S. population (who are more likely to have mixed ethnic ancestries) makes population stratification an unlikely explanation for the results in these case-control studies [10–12].

Our finding of a lack of association between the *COMT* codon 158 variant A allele and breast cancer is similar to the two prior case-control studies conducted among Scandinavian populations (one in Sweden and another in Finland) where the possibility of population stratification bias is less likely (since genetic background of the cases and controls were likely to be similar) [22, 23]. The case-control studies that showed positive associations were conducted in the US and the Asian populations in which the possibility of population stratification is likely. However, since our study population is of modest size, the statistical stability of our findings may also be questioned.

Our study also evaluated whether parental genotypes are associated with the breast cancer risk in the daughters (independent of daughter's own genotype) in addition to evaluating main associations between a woman's breast cancer risk and her own constitutional genotype. We found that parental (both maternal and paternal) carrier status for the *CYP1B1* codon 453 G allele and the *COMT* codon 158 variant A allele is associated with breast cancer risk in daughters. The association in relation to maternal carrier status is biologically relevant to evidence from the prior literature on the association between *in utero* exposure to hormonal factors and breast cancer risk in adulthood [24–29]. However, the association in relation to paternal carrier status is an unexpected finding and is not consistent with the *in utero* hormonal exposure hypothesis. While these novel associations, if real, may prove to have important implications in advancing our understanding of the breast cancer etiology, given that this is the first study showing such associations and also in the face of lack of any main association, we believe that future studies need to confirm them before any meaningful inferences can be made.

Several limitations of our study warrant caution for interpreting its findings. First, the study included nuclear families that were not selected based on a population-based scheme. While this may limit the generalizability of the findings, it should not affect the validity of the observed (or lack of) associations. Second, as mentioned above, the study sample is of modest size and thus the possibility of failing to obtain a statistically significant association due to lack of adequate power may not be ruled out. We may not have had enough statistical power to detect small increases in risk associated with certain of the *CYP1B1* or *COMT* genotypes. In particular, for discordant sib-ship analyses, we lacked power to evaluate interactions between genotypes and both endogenous and exogenous hormonal characteristics, such as age at menarche, timing and number of pregnancies and use of exogenous hormones. The evaluation of such interactions will be the subject of separate future analyses, based on additional numbers of CFR families. Third, the current study employed multiple comparisons: three different alleles for two statistics under three different genetic models leading to 18 comparisons. It is possible that multiple comparisons may have led to one or more of the observed associations, especially those relating parental carrier status and risk in offspring. Fourth, the fact that these associations were observed for both paternal and maternal carrier status, may be suggestive against a biological specificity of the finding. Finally, the functions of the examined variants in the *CYP1B1* and *COMT* genes are not clearly known. Moreover, the CYP1B1 gene is known to be inducible, which makes the functional significance of the variants more difficult to interpret.

In conclusion, our family-based study found that the *CYP1B1* codon 432 variant C and codon 453 variant G alleles and *COMT* codon 158 variant A allele are not associated with breast cancer among families participating in the MNYR breast cancer family registry. Although the study suggests that parental carrier status of these alleles is associated with breast cancer in daughters in these families, these associations should be interpreted with caution and need to be confirmed in future larger family-based studies.

Acknowledgements

This research was supported by U.S. Department of Defense Grants # DAMD17-00-1-0213 and # DAMD17-02-1-0354 and National Cancer Institute under RFA # CA-95-003, Breast Cancer Family Registry, through cooperative agreement with Columbia University.

References

1. Service RF: New role for estrogen in cancer? Science 279: 1631–1633, 1998
2. Jan ST, Devanesan PD, Stack DE, Ramanathan R, Byun J, Gross ML, Rogan EG, Cavalieri EL: Metabolic activation and formation of DNA adducts of hexestrol, a synthetic nonsteroidal carcinogenic estrogen. Chem Res Toxicol 11: 412–419, 1998
3. Sutter TR, Tang YM, Hayes CL, Wo YY, Jabs EW, Li X, Yin H, Cody CW, Greenlee WF: Complete cDNA sequence of a human dioxin-inducible mRNA identifies a new gene subfamily of cytochrome P450 that maps to chromosome 2. J Biol Chem 269: 13092–13099, 1994
4. McKay JA, Melvin WT, Ah-See AK, Ewen SW, Greenlee WF, Marcus CB, Burke MD, Murray GI: Expression of cytochrome P450 CYP1B1 in breast cancer. FEBS Lett 374: 270–272, 1995
5. Taylor MC, McKay JA, Murray GI, Greenlee WF, Marcus CB, Burke MD, Melvin WT: Cytochrome P450 1B1 expression in human malignant tumours. Biochem Soc Trans 24: 328S, 1996
6. Spink DC, Spink BC, Cao JQ, DePasquale JA, Pentecost BT, Fasco MJ, Li Y, Sutter TR: Differential expression of CYP1A1 and CYP1B1 in human breast epithelial cells and breast tumor cells. Carcinogenesis 19: 291–298, 1998
7. Murray GI, Taylor MC, McFadyen MC, McKay JA, Greenlee WF, Burke MD, Melvin WT: Tumor-specific expression of cytochrome P450 CYP1B1. Cancer Res 57: 3026–3031, 1997
8. Tang YM, Wo YY, Stewart J, Hawkins AL, Griffin CA, Sutter TR, Greenlee WF: Isolation and characterization of the human cytochrome P450 CYP1B1 gene. J Biol Chem 271: 28324–28330, 1996
9. Bailey LR, Roodi N, Dupont WD, Parl FF: Association of cytochrome P450 1B1 (CYP1B1) polymorphism with steroid receptor status in breast cancer. Cancer Res 58: 5038–5041, 1998
10. Zheng W, Xie DW, Jin F, Cheng JR, Dai Q, Wen WQ, Shu XO, Gao YT: Genetic polymorphism of cytochrome P450-1B1 and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 9: 147–150, 2000
11. Watanabe J, Shimada T, Gillam EM, Ikuta T, Suemasu K, Higashi Y, Gotoh O, Kawajiri K: Association of CYP1B1 genetic polymorphism with incidence to breast and lung cancer. Pharmacogenetics 10: 25–33, 2000
12. Kocabas NA, Sardas S, Cholerton S, Daly AK, Karakaya AE: Cytochrome P450 CYP1B1 and catechol *O*-methyltransferase (COMT) genetic polymorphisms and breast cancer susceptibility in a Turkish population. Arch Toxicol 76: 643–649, 2002
13. Lotta T, Vidgren J, Tilgmann C, Ulmanen I, Melen K, Julkunen I, Taskinen J: Kinetics of human soluble and membrane-bound catechol *O*-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme. Biochemistry 34: 4202–4210, 1995
14. Lavigne JA, Helzlsouer KJ, Huang HY, Strickland PT, Bell DA, Selmin O, Watson MA, Hoffman S, Comstock GW, Yager JD: An association between the allele coding for a low activity variant of catechol-*O*-methyltransferase and the risk for breast cancer. Cancer Res 57: 5493–5497, 1997
15. Thompson PA, Shields PG, Freudenheim JL, Stone A, Vena JE, Marshall JR, Graham S, Laughlin R, Nemoto T, Kadlubar FF, Ambrosone CB: Genetic polymorphisms in catechol-*O*-methyltransferase, menopausal status, and breast cancer risk. Cancer Res 58: 2107–2110, 1998
16. Millikan RC, Pittman GS, Tse CK, Duell E, Newman B, Savitz D, Moorman PG, Boissy RJ, Bell DA: Catechol-*O*-methyltransferase and breast cancer risk. Carcinogenesis 19: 1943–1947, 1998
17. Goodman JE, Lavigne JA, Wu K, Helzlsouer KJ, Strickland PT, Selhub J, Yager JD: COMT genotype, micronutrients in the folate metabolic pathway and breast cancer risk. Carcinogenesis 22: 1661–1665, 2001
18. Huang CS, Chern HD, Chang KJ, Cheng CW, Hsu SM, Shen CY: Breast cancer risk associated with genotype polymorphism of the estrogen-metabolizing genes CYP17, CYP1A1, and COMT: a multigenic study on cancer susceptibility. Cancer Res 59: 4870–4875, 1999
19. Yim DS, Parkb SK, Yoo KY, Yoon KS, Chung HH, Kang HL, Ahn SH, Noh DY, Choe KJ, Jang IJ, Shin SG, Strickland PT, Hirvonen A, Kang D: Relationship between the Val158Met polymorphism of catechol *O*-methyl transferase and breast cancer. Pharmacogenetics 11: 279–286, 2001
20. Matsui A, Ikeda T, Enomoto K, Nakashima H, Omae K, Watanabe M, Hibi T, Kitajima M: Progression of human breast cancers to the metastatic state is linked to genotypes of catechol-*O*-methyltransferase. Cancer Lett 150: 23–31, 2000
21. Hamajima N, Matsuo K, Tajima K, Mizutani M, Iwata H, Iwase T, Miura S, Oya H, Obata Y: Limited association between a catechol-*O*-methyltransferase (COMT) polymorphism and breast cancer risk in Japan. Int J Clin Oncol 6: 13–18, 2001
22. Mitrunen K, Jourenkova N, Kataja V, Eskelinen M, Kosma VM, Benhamou S, Kang D, Vainio H, Uusitupa M, Hirvonen A: Polymorphic catechol-*O*-methyltransferase gene and breast cancer risk. Cancer Epidemiol Biomarkers Prev 10: 635–640, 2001
23. Bergman-Jungestrom M, Wingren S: Catechol-*O*-Methyltransferase (COMT) gene polymorphism and breast cancer risk in young women. Br J Cancer 85: 859–862, 2001
24. Sanderson M, Williams MA, Daling JR, Holt VL, Malone KE, Self SG, Moore DE: Maternal factors and breast cancer risk among young women. Paediatr Perinat Epidemiol 12: 397–407, 1998
25. Weiss HA, Potischman NA, Brinton LA, Brogan D, Coates RJ, Gammon MD, Malone KE, Schoenberg JB: Prenatal and perinatal risk factors for breast cancer in young women. Epidemiology 8: 181–187, 1997
26. Ekbom A, Hsieh CC, Lipworth L, Adami HQ, Trichopoulos D: Intrauterine environment and breast cancer risk in women: a population-based study. J Natl Cancer Inst 89: 71–76, 1997
27. Potischman N, Troisi R: *In-utero* and early life exposures in relation to risk of breast cancer. Cancer Causes Control 10: 561–573, 1999

28. Ekbom A, Trichopoulos D, Adami HO, Hsieh CC, Lan SJ: Evidence of prenatal influences on breast cancer risk. Lancet 340: 1015–1018, 1992
29. Sanderson M, Williams MA, Malone KE, Stanford JL, Emanuel I, White E, Daling JR: Perinatal factors and risk of breast cancer. Epidemiology 7: 34–37, 1996
30. Ahsan H, Chen Y, Whittemore AS, Senie RT, Hamilton SP, Wang Q, Gurvich I, Santella RM: Variants in estrogen-biosynthesis genes CYP17 and CYP19 and breast cancer risk: a family-based genetic association study (submitted for publication)
31. Hsu TM, Chen X, Duan S, Miller RD, Kwok PY: Universal SNP genotyping assay with fluorescence polarization detection. Biotechniques 31: 560, 562, 564–560,8, passim, 2001
32. Chen X, Levine L, Kwok PY: Fluorescence polarization in homogeneous nucleic acid analysis. Genome Res 9: 492–498, 1999
33. Tu IP, Balise RR, Whittemore AS: Detection of disease genes by use of family data. II. Application to nuclear families. Am J Hum Genet 66: 1341–1350, 2000
34. Whittemore AS, Tu IP: Detection of disease genes by use of family data. I. Likelihood-based theory. Am J Hum Genet 66: 1328–1340, 2000
35. Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52: 506–516, 1993
36. Schaid DJ: General score tests for associations of genetic markers with disease using cases and their parents. Genet Epidemiol 13: 423–449, 1996
37. Spielman RS, Ewens WJ: The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet 59: 983–989, 1996
38. Hosmer D, Lemeshow S: Logistic Regression for Matched Case-Control Studies. Applied Logistic Regression, John Wiley & Sons, New York, 1989, pp 187–213
39. Breslow NE, Day NE: Statistical Methods in Cancer Research. – The Analysis of Case-Control Studies. Vol. I, IARC Sci. Publ. 1980, pp. 5–338
40. De VI, Hankinson SE, Li L, Colditz GA, Hunter DJ: Association of CYP1B1 polymorphisms and breast cancer risk. Cancer Epidemiol Biomarkers Prev 11: 489–492, 2002

Address for offprints and correspondence: Dr Habibul Ahsan, Department of Epidemiology, Mailman School of Public Health, 722 West 168th Street, Room 720G, New York, NY 10032, USA; *Tel.:* +1-212-305-4657; *Fax:* +1-212-305-9413; *E-mail:* ha37@columbia.edu